Optik 126 (2015) 1269–1276

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

# A multi-scale method for urban tree canopy clustering recognition using high-resolution image

Jian-Nong Cao[a], Zhenfeng Shao[b,*], Jia Guo[a], Bei Wang[a], Yuwei Dong[a], Pinglu Wang[a]

[a] The School of Earth Science and Resources, Chang'an University, No. 126, Yanta South Road, Xi'an 710054, China
[b] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, 129 Luoyu Road, Wuhan 430079, China

ARTICLE INFO

*Article history:*
Received 10 February 2014
Accepted 18 February 2015

*Keywords:*
Image segmentation
High-resolution image
Constraints mean shift
Wavelet
Tree canopy

ABSTRACT

In this paper a constraint mean shift method is proposed for extracting urban tree canopy with the use of high-resolution image. Through establishing multi-scale pyramid image feature space by wavelet, features between layers can be combined as the constrained route for mean shift to realize self-adaptive decomposition and scale transfer in multi-scale feature space, then differences of internal and external structure in urban tree canopy and differences of average spectral radiant intensity are used as multi-scale feature space of wavelet to realize the preliminary clustering segmentation, finally we apply the supervised segmentation to extract tree canopy based on clustering feature. Experiments demonstrate that the proposed method can eliminate the over-detailed effect of image accurate extraction of the urban tree canopy can be achieved.



## 1. Introduction

Urban vegetation coverage has a great impact on the environment and its accurate estimation is of significance to guide the future urban planning management and environmental protection. Urban tree canopy is an important part of eco-system and an important biophysical index for ecological evaluation. Living vegetation volume (LVV) is defined as the volume of space occupied by stems and leaves. It is also the first 3-dimensional index of the urban greening index system. LVV has great significance in the research of greening layout, quantity and their correlation with atmospheric environment [1]. With the rapid change of urban environment, there is an urgent need for using the remote sensing technique to calculate large area coverage.

The aerial remote sensing, multi-spectral remote sensing, high-resolution remote sensing images are the main data sources for information extraction of urban tree canopy. Conventional supervised and unsupervised classification is the primary method to extract the urban tree canopy information [2]. Literatures [3–5] studied on extracting the information of tree canopy by high-resolution sensing images using methods including learning and supervision, region-based growing, watershed, template matching and the mixture of these methods, but few consideration has been focused on the role of image scale on tree canopy recognition.

Literature [6] presented a multi-scale texture segmentation method whose multi-scale pyramid feature space was based on the mean of low-frequency and standard variances of high-frequency decomposed by the wavelet transform coefficient, applying mean shift algorithm to decompose the feature space, separating between different areas in coarse scale, locating the edge and other details in finer scale, and transferring the clustering results from coarse to finer scale layer-by-layer with a proper strategy, which led to a successful multi-scale texture segmentation. However, this method destroyed the smoothness of the feature space at different scales during transferring multi-scale features and increased the difficulty of parameters selecting, such as the bandwidth of mean shift (MS).

Thus, this paper integrates the spectral information (the mean of wavelet transform low-frequency coefficient) and texture information (the standard variances of wavelet transform high-frequency coefficient) in multi-scale structures (wavelet transform pyramid structure). Pyramid structure of the multi-scale feature space is established by wavelet decomposition. Image features are calculated by differences of internal and external urban tree canopy structure and difference of the average spectral radiant intensity, presenting the constraints mean shift (CMS) method to solve the complexity of feature space caused by multi-scale clustering transfer, and overcome the defects of target recognition caused by the lack of basic feature factors (spectrum, texture or scale). Self-adaptive decomposition of feature space is conducted with the

* Corresponding author. Tel.: +86 15827188114.
*E-mail addresses:* caojiannong@126.com (J.-N. Cao), shaozhenfeng@whu.edu.cn (Z. Shao).

http://dx.doi.org/10.1016/j.ijleo.2015.02.094
0030-4026/© 2015 Published by Elsevier GmbH.

constraints mean shift (CMS) algorithm, which reduces the resolution of the image segmentation. Moreover, this paper presents a supervised segmentation method based on clustering feature, which further refines the segmentation results.

The proposed method works better on segmentation of high-resolution image. The rest of this paper is organized as follows: the first section discusses the constraints mean shift clustering; the second section is the extraction algorithm of the urban tree canopy based on clustering features; the third section is the experiment and analysis; the fourth section is the conclusion.

## 2. Constraints mean shift clustering

Wavelet transformation provides a comprehensive method texture analysis, which integrated of frequency spectrum, image structure and pixel statistics. Through decomposing remote sensing image via computational functions involving flexible, shift, etc., wavelet transformation holds strong capability in spatial and frequency decomposition. As an image analysis tool with high performance, wavelet transformation could give unified framework for image texture features representation and extraction. In this study, low pass image signal is recursive decomposition, each scale features are combined with low pass and high pass signals, respectively. Texture features of image regions would be discomposed in the successive layer, the pyramid structure of wavelet coefficients feature space can be constructed by salient feature regions in each scales.

$$W_f(a,b)=\frac{1}{\sqrt{|a|}}\int_{-\infty}^{+\infty} f(t)\psi\left(\frac{t-b}{a}\right)dt \quad (1)$$

where $\psi(x)\in L^2(R)$ is wavelet basis, $a$ is flexible factor, $b$ is shift factor. If binary discretization is operated for scale factor of continued wavelet transformation and shift factor ($a=2^j$, $b=2^jk$, $(j,k\in Z)$), there would be produced a wavelet with scale and time discretization.

Some literatures demonstrated the convergence [7,8] and sufficient condition of convergence [9] from mean shift algorithm on continuous and discrete function. Literature [9] successfully applied the algorithm on image segmentation and smoothing. It obtained the extensive application on image filtering, edge extraction, information fusion and other fields of image processing.

The Eq. (2) is applied to conduct iterative computations in literature [9].

$$y_{j+1}=\frac{\sum_{i=1}^{n} x_i g\left(\left\|\frac{y_i-x_i}{h}\right\|^2\right)}{\sum_{i=1}^{n} g\left(\left\|\frac{y_i-x_i}{h}\right\|^2\right)},\quad j=1,2,\ldots \quad (2)$$

where $y_{j+1}$ is the process feature after the $j$ time iteration, in Eq. (1), $x$ is the current point, the process feature which is the closest to $y_{j+1}$, namely data center within bandwidth. Literature [9] demonstrated that sequence $y_{j+1}$ can do convergence to $y'$ when meeting certain conditions, then $y'$ can be called corresponding pattern of $x$. The complexity of mean shift technique relies on two correlative controlled parameter, namely bandwidth $h$ and density $T$ [8].

In image processing, except for the feature space of transformation, the spatial information within every pixel is also important. So, it is necessary to apply the spatial information of image to feature vectors, then, the kernel function is changed as following:

$$K_{h_s^2,h_r^p}(x)=\frac{C}{h_s^2h_r^p}k\left(\left\|\frac{x^s}{h_s}\right\|^2\right)k\left(\left\|\frac{x^r}{h_r}\right\|^2\right) \quad (3)$$

where $x^r$ is the feature of space location, $x^s$ is feature vector that constitutes feature space, $h_r$ and $h_s$ is separately space-bandwidth and feature-bandwidth, value of the window about space-width in image is $h_r=(2n_r+1)\times(2n_r+1)$, $C$ is normalization constant.

The density surface of discrete feature space in factual image is normally unsmooth, and density statistics is closely related to the selection of bandwidth in non-uniform discrete data [8–10]. If the bandwidth is too large, it may cause over-generalized density pattern distortion; if too small, it may easily cause density islands and other over-trivial density patterns in sparse data area, making the algorithm fall into local maximum. So, it is a reliable method [8] to utilize smaller bandwidth with attached density constraint. Density constraint cannot only avoid insignificance pattern's judgement procedure, but also improve clustering accuracy and speed up the convergence.

The difficulty of bandwidth selection is that both the narrow peak of statistical density model and the variation of details are related to bandwidth, literature [11] proposed two self-adaptive schemes: one is non-parametric method, which needs to define new self-adaption process of mean shift, applying insert rule and sampling point density estimation operator; the other is semi-parametric method, applying local data structure to extract scale information. Selection of clustering in self-adaptive bandwidth is superior to the fixed bandwidth method. However, it also increases the computation cost of mean shift, on the other hand. In addition complexity of image feature space is hard to be adapted by insert rule and structure information. Generally speaking, the peak-to-valley ratio method proposed by literature [8] is more reliable and practical, but the computation of peak-to-valley ratio needs to build density surface model which increases extra computing burden. Based on the above analysis, this paper proposed the constraints mean shift (CMS) method to solve the robustness of mean shift in complex feature space, bandwidth constraint information is from merging features transferred by scale, with no extra computing.

From Eq. (2), we can find that the conventional mean shift algorithm assumes that density smoothing is given with a limit of density change rate. As indicated in Fig. 1, the blue mass points can be used for obtaining a smooth feature space, in this way the algorithm can be successfully implemented. However, the red mass points in Fig. 1 are fragmented, discontinuous, stair-step unsmooth, indicating the difficulty to obtain a smooth feature space, so the reasonability and continuity of the mean shift route is invalid with Eqs. (1) and (2), and the algorithm is difficult to implement in this kind of feature space. Studying on the constraints mean shift (CMS), it

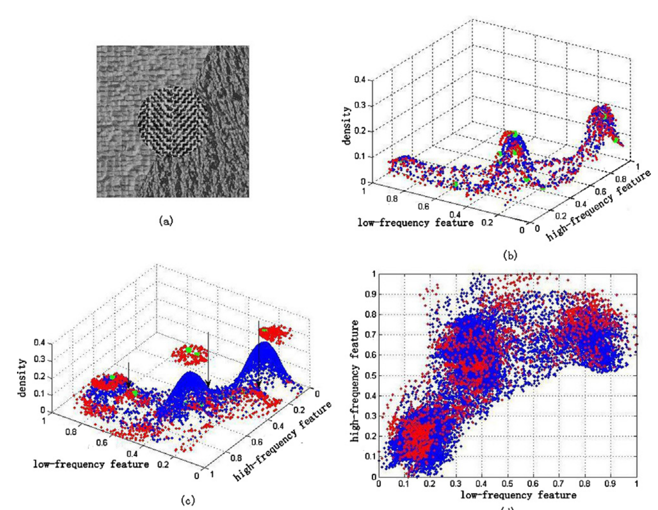


**Fig. 1.** Feature density distribution of scale feature transfer plot: (a) three composite image; (b) the density distribution of the fourth layer on the third layer; (c) the density distribution of the second layer on the first layer; (d) feature density distribution of the second layer on the first layer.